S-arrangements avec répétitions de tas

Sylviane R. Schwer

S-arrangements avec répétitions de tas

Sylviane R. Schwer
LIPN, URPR-A CNRS 7030, Université Paris 13, Institut Galilée
99, Avenue Jean-Baptiste Clément 93430 Villetaneuse, France
fax : +33 (1) 48.26.07.12 tél. : +33 (1) 49.40.36.84
e-mail : schwer@lipn.univ-paris13.fr

20 mars 2008

Résumé

Nous étudions quelques aspects combinatoires de l'ensemble des arrangements de blocs avec répétitions de tas ou multi-ensembles en utilisant le cadre des langages formels.

mots clefs : combinatoire énumérative, langages formels.

1 Introduction

Donnons-nous n couleurs différentes : $(c_1, \ldots, c_n)$, et pour chaque $i \in [n]$, p_i boules indiscernables de couleur c_i ainsi que r boîtes alignées, avec $max(p_1, \ldots, p_n) \leq r \leq p_1 + \cdots + p_n$. Nous proposons d'examiner le problème suivant : placer les boules dans les boîtes de telle façon que toutes les boîtes soient occupées et qu'il n'y ait pas deux boules de même couleur dans la même boîte. Le cas où tous les p_i valent 1 a été étudié par [5] sous le terme d'arrangement préférentiel ; le cas où tous les p_i valent 2 a été étudié en Intelligence Artificielle [3]. Rappelons, suivant [2, 25–28], que si X est un n-ensemble, un bloc de X – ou X-bloc – est une partie non vide de X et un tas sur X – ou X-tas – ν est une application de X dans $\mathbb{N}$.

Définition 1.1 *Soit ν un tas sur un n-ensemble X, un S-arrangement avec répétitions de ν est une liste de X-blocs, les répétitions étant autorisées, chaque élément $x \in X$ apparaissant dans exactement $\nu(x)$ X-blocs. On note $\mathfrak{D}(\vec{\nu}(X))$* [resp. $\mathfrak{D}_k(\vec{\nu}(X))$] *l'ensemble des S-arrangements* [resp. *de longueur k*] *du X-tas ν, et $D(\vec{\nu}(X))$* [resp.$D_k(\vec{\nu}(X))$] *le nombre de S-arrangements du X-tas ν* [resp. *de longueur k*].

Le formalisme de la théorie des langages est bien adapté à ce problème [2], c'est pourquoi nous définissons dans ce cadre les outils nécessaires. Considérant $X = (x_1, \ldots, x_n)$ comme un alphabet ordonné, on représente ν comme le mot $x_1^{\nu(x_1)} \ldots x_n^{\nu(x_n)}$ sur X^*. Le vecteur $\vec{\nu}(X) = (\nu(x_1), \ldots, \nu(x_n))$ est appelé vecteur de Parikh [4] plutôt que spécification [2]. Si $\vec{\nu}(X) = (p_1, \ldots, p_n)$, alors on note $\mathfrak{D}_X(p_1, \ldots, p_n)$ et $D_X(p_1, \ldots, p_n)$ à la place de $\mathfrak{D}(\vec{\nu}(X))$ et $D(\vec{\nu}(X))$.

2 Les mots pour le dire

Nous utilisons le vocabulaire classique de la théorie des langages [1, 4]. Le mot vide est noté ε, $|f|$ est la longueur du mot f, $|f|_x$ le nombre d'occurrences de la lettre x dans l'écriture de f. L'entrelacs (shuffle [4, p.108]) de deux mots f et g de X^* est le langage $f \Cup g = \{f_1 g_1 \dots f_n g_n \| \forall i : f_i, g_i \in X^*,\ f_1 \dots f_n = f,\ g_1 \dots g_n = g\}$. Ainsi $aa \Cup bb = \{aabb, abab, baab, abba, baba, bbaa\}$. L'entrelacs de deux langages L et L' est le langage $L \Cup L' = \{f \Cup f' | f \in L, f \in L'\}$. L'entrelacs munit $\wp(X^*)$, ensemble des langages sur X, d'une structure de monoïde commutatif dont l'élément neutre est l'ensemble $\{\varepsilon\}$. Un résultat classique de combinatoire sur les mots est que $card(f \Cup g) = \binom{|f+g|}{|f|}$ pour peu que f et g n'aient pas de lettre commune. On en déduit par récurrence le

Lemme 2.1 *Soit $X^n = \{x_1, \dots, x_n\}$ un n-alphabet, $p_1, \dots, p_n$ n entiers positifs,*

$$card(x_1^{p_1} \Cup \cdots \Cup x_n^{p_n}) = \frac{(p_1 + \cdots + p_n)!}{p_1! \cdots p_n!} = \binom{p_1 + \cdots + p_r}{p_1, \cdots, p_r}$$

Ce nombre est connu comme un coefficient multinômial.

Définition 2.1 (S-alphabets et S-langages) *Soit X un alphabet, un* S-alphabet sur X *est une partie non vide de 2^X. Un élément d'un S-alphabet sur X est une* S-lettre*. Un* S-mot *sur X est une suite finie de S-lettres sur X, un ensemble de S-mots sur un S-alphabet est un* S-langage.

Nous allons utiliser les S-alphabets sur X suivants : (i) le S-alphabet *naturel* $\dot{X} = \{\{a\} | a \in X\}$ que l'on identifie à X et (ii) le S-alphabet *saturé* sur X, *i.e.* $\widehat{X} = 2^X - \emptyset$.

On a donc $X \subset \widehat{X}$ et $X^* \subset \widehat{X}^*$. On écrira les S-lettres aussi bien horizontalement que verticalement. Ainsi $\widehat{\{a, b\}} = \{a, b, \left\{ {a \atop b} \right\}\}$.

Les S-alphabets étant des alphabets, on peut donc travailler avec les outils connus de la théorie des langages. Néanmoins, il nous faut introduire des notations pour relier les S-mots à l'alphabet de base et non à leur S-alphabet.

Définition 2.2 *Soit $X = \{x_1, \dots, x_n\}$ un n-alphabet et $f \in \widehat{X}^*$. On note $\|f\|_x$ le nombre d'occurrences de la lettre $x \in X$ dans le S-mot f. On appelle* taille *de f et on note $\|f\|$ l'entier $\sum_{1 \le i \le n} \|f\|_{x_i}$.*

La définition qui suit généralise celle de [4, p.146].

Définition 2.3 *Soit $X = \{x_1, \dots, x_n\}$ un n-alphabet et $f \in \widehat{X}^*$. On appelle* vecteur de Parikh *de f, noté $\vec{f}$, le n-uplet $(\|f\|_{x_1}, \dots, \|f\|_{x_n})$.*

Ainsi, sur l'alphabet $X = \{a, b\}$, le S-mot $f = ab \left\{ {a \atop b} \right\} bab$ a pour longueur 6, pour taille 7, et pour vecteur de Parikh $\vec{f} = (3, 4)$. Les occurrences de a occupent les rangs 1, 3, 5 ; les occurences de b occupent les rangs 2, 3, 4 et 6. f représente un k-S-arrangement avec répétition du tas $\{a^3, b^4\}$.

Nous étendons également l'opération d'entrelacs aux S-mots et S-langages en autorisant le partage de S-lettres. Nous donnons ici une définition restreinte au cas où les alphabets sont disjoints.

Définition 2.4 *Soit X et Y deux alphabets disjoints, $f \in \widehat{X}^*$, $g \in \widehat{Y}^*$ le* S-entrelacs *de f et g est le langage $f \mathbin{\overline{\Cup}} g = \{h_1 \dots h_r \mid h_i \in \widehat{X \cup Y},\ 0 \le |f_i|, |g_i| \le 1,\ max(|f|, |g|) \le r \le |f| + |g|,\ f = f_1 \dots f_n,\ g = g_1 \dots g_n,$ et $\forall i \in [r],\ 1 \le |f_i| + |g_i| \le 2, h_i = f_i \cup g_i\ \}$.*

Ainsi $aa \mathbin{\overline{\Cup}} bb$=\{*aabb*, *a* $\left\{ {a \atop b} \right\}$ *b*, *abab*, $\left\{ {a \atop b} \right\}$ *ab* ,*ab* $\left\{ {a \atop b} \right\}$, $\left\{ {a \atop b} \right\} \left\{ {a \atop b} \right\}$, *baab*, *ba* $\left\{ {a \atop b} \right\}$, *abba*, $\left\{ {a \atop b} \right\}$ *ba*, *baba*, *b* $\left\{ {a \atop b} \right\}$ *a*, *bbaa*\}.

Le S-entrelacs de deux S-langages L et L' écrits sur des alphabets disjoints est le langage $L \mathbin{\overline{\Cup}} L' = \cup_{f \in L, f' \in L'} f \mathbin{\overline{\Cup}} f'$. Cette opération est associative, commutative et possède comme élément

neutre $\{\varepsilon\}$.
Si le n-tas ν sur l'alphabet $X = \{x_1, \ldots, x_n\}$ vaut $\{x_1^{p_1}, \ldots, x_n^{p_n}\}$, on écrit $\mathbb{W}\nu_X = x_1^{p_1}\mathbb{W}\ldots\mathbb{W}x_n^{p_n}$.

Le résultat primordial est le suivant :

Théorème 2.1 *Soit ν un n-tas sur l'alphabet X. Il existe des bijections naturelles entre : (i) $\{f \in \widehat{X}^* | \vec{f} = \vec{\nu}(X)\}$, (ii) $\mathfrak{D}(\nu(X))$ et (iii) $\mathbb{W}\nu_X$.*

Preuve : Soit $\nu = \{x_1^{p_1}, \ldots, x_n^{p_n}\}$ un n-tas sur l'alphabet $X = \{x_1, \ldots, x_n\}$, on a les trois ensembles (i) $\{f \in \widehat{X}^* | \vec{f} = (p_1, \ldots, p_n)\}$, (ii) $\mathfrak{D}_X(p_1, \ldots, p_n)$ et (iii) $x_1^{p_1}\mathbb{W}\ldots\mathbb{W}x_n^{p_n}$. Chaque bloc de X correspond à une S-lettre sur X, d'où la bijection entre (i) et (ii). Tout mot de (iii) est un mot de (i) ; réciproquement tout S-mot de (i) possède exactement p_i S-lettres contenant chacune une et une seule occurrence de x_i, la projection d'un tel S-mot sur $\{x_i\}$ est bien x^{p_i}. $\square$

3 Enumérations

Classiquement, on partitionne un langage soit en classant ses mots en fonction de leur lettre d'une extrémité, soit en fonction de leur longueur, soit encore suivant une décomposition pertinente. Partitionner $\mathbb{W}\nu_X$ suivant la S-lettre extrême de ses S-mots donne une fonction récursive dont la fonction génératrice à $card(X)$ variables est simple. Le partitionner en fonction de la longueur des S-mots nous fournit une somme de coefficients multinomiaux. Exprimer $\mathbb{W}\nu_X$ de façon itérative fournit une décomposition de ses S-mots en fonction d'une lettre particulière et nous donne encore une autre énumération.

Soit $(p_1, \ldots, p_n)$ n entiers non négatifs, posons $\mathcal{P}red(p_1, \ldots, p_n) = \{(\widetilde{p_1}, \cdots, \widetilde{p_n}) \in \mathbb{N}^n | \ si\ p_i \neq 0$, alors $\widetilde{p_i} \in \{p_i, p_i - 1\}$ sinon $\widetilde{p_i} = p_i,\ 1 \le i \le n\} - \{(p_1, \cdots, p_n)\}$. Le classement suivant la dernière S-lettre donne :

Proposition 3.1 *Soit $\nu_X = \{x_1^{p_1}, \ldots, x_n^{p_n}\}$ un n-tas.*

$$\mathbb{W}\nu_X = \bigcup_{\mathfrak{a} \in \widehat{X}} \mathbb{W}\{x_1^{p_1 - \|\mathfrak{a}\|_{x_1}}, \ldots, x_n^{p_n - \|\mathfrak{a}\|_{x_n}}\}.\mathfrak{a}$$

et

$$D(p_1, \ldots, p_n) = \sum_{\mathcal{P}red(p_1, \ldots, p_n)} D_n(\widetilde{p_1}, \cdots, \widetilde{p_n})$$

avec $D_1(0) = 1,\ D_n(p_1, \ldots, p_{n-1}, 0) = D_{n-1}(p_1, \ldots, p_{n-1})$.

La fonction génératrice sur n variables $(X_1, \ldots, X_n)$ vaut :

$$\sum_{p_1, \cdots, p_n \ge 0} D_n(p_1, \cdots, p_n) X_1^{p_1} \ldots X_n^{p_n} = \frac{1}{2 - \prod_{1 \le i \le n}(1 + X_i)} = \frac{1}{1 - \sum_{\emptyset \neq s \subseteq [1,n]} \prod_{i \in s} X_i}$$

$D_n(p_1, \cdots, p_n)$ est une fonction symétrique. Ainsi, on peut supprimer tous les p_i nuls et les dimensions correspondantes. Dans le cas où tous les p_i sont égaux à p on a

$$D(\underbrace{p, \ldots, p}_{n\ fois}) = \sum_{r=1}^{n} \binom{n}{r} D(\underbrace{p-1, \ldots, p-1}_{r\ fois}, \underbrace{p, \ldots, p}_{n-r\ fois})$$

Les nombres $D(p_1, p_2)$ sont les nombres de Delannoy [2, p.93].

Maintenant, et si $\mathbb{P}$ désigne l'ensemble des entiers positifs, notons $\mathcal{V}_r(p_1,\ldots,p_n)$ l'ensemble $\{(t_1S_1,\ldots,t_rS_r)|\ \ (t_1,\ldots,t_k)\in\mathbb{P}^r,\ \ (S_1,\ldots,S_k)\in\widehat{X}^r$, tels que $\forall i\in[n],\ \sum_{k=1}^{r}t_k\|S_k\|_{x_i}=p_i\}$. En utilisant le lemme 2.1, le classement suivant la longueur des S-mots donne :

Proposition 3.2 *Soit* $\nu_X=\{x_1^{p_1},\ldots,x_n^{p_n}\}$ *un n-tas.*

$$\mathbb{W}\nu_X=\bigcup_{r=max(p_1,\ldots,p_n)}^{p_1+\cdots+p_n}\ \bigcup_{\mathcal{V}_r(p_1,\ldots,p_n)}S_1^{p_1}\mathbb{W}\cdots\mathbb{W}S_r^{p_r}$$

et

$$D(p_1,\ldots,p_n)=\sum_{r=max(p_1,\ldots,p_n)}^{p_1+\cdots+p_n}\ \sum_{\mathcal{V}_r(p_1,\ldots,p_n)}\frac{(t_1+\cdots+t_k)!}{t_1!\cdots t_k!}$$

Enfin, l'énumération par récurrence sur le cardinal de l'alphabet donne :

Proposition 3.3 *Soit* $\nu_X=\{x_1^{p_1},\ldots,x_n^{p_n}\}$ *un n-tas.*

$$\mathbb{W}\nu_X=\mathbb{W}(x_1^{p_1},\ldots,x_{n-1}^{p_{n-1}})\mathbb{W}x_n^{p_n}$$

et

$$D(p_1,\ldots,p_n)=\sum_{r=max(p_1,\ldots,p_n)}^{p_1+\cdots+p_n}\lambda_{\vec{\nu}(X)}(r,n)$$

avec $\lambda_{\vec{\nu}(X)}(r,n)=\binom{r}{p_n}(\sum_{k=0}^{p_n}\binom{p_n}{k}\lambda_{\vec{\nu}(X)}(r-k,n-1))$ *et* $\lambda_{\vec{\nu}(X)}(r,1)=1$ *si* $r=p_n$ *et* 0 *sinon, en notant* $\lambda_{\vec{\nu}(X)}(r,n)$ *le nombre de S-mots de* ν_X *de longueur* r*.*

Cela se fait en distinguant les p_n S-lettres contenant un x_n. Ces S-lettres peuvent être des singletons (donc disparaître par élimination des x_n) ou non (donc rester).

4 Conclusion et perspectives

Nous avons présenté une famille d'objets appelés S-arrangements avec répétions de n-tas et défini les outils de théorie des langages associés. Quelques résultats combinatoires ont été donnés. Cette famille d'objets généralise plusieurs familles d'objets déjà existants en combinatoire et en informatique théorique. Il y a beaucoup d'autres familles d'objets que l'on peut représenter en ces termes par une bijection naturelle. Ceci fera l'objet de publications ultérieures. D'autre part, nous étudions les propriétés de treillis et de génération par systèmes de réécriture du langage $x_1^{p_1}\mathbb{W}\ldots\mathbb{W}x_n^{p_n}$.

Références

[1] J.-M. Autebert. *Langages Algébriques.* Masson 1987

[2] L. Comtet. *Analyse Combinatoire. Tome premier.* P.U.F. Collection SUP, 1970

[3] M. Dubois et S. R. Schwer. Classification topologique des ensembles convexes de Allen. Proceedings de R.F.I.A. 2000, Paris, p.59-68

[4] S. Ginsburg. *The Mathematical Theory of Context-Free Languages.* McGraw-Hill 1966

[5] O. A. Gross. Preferential arrangements. Amer. Math. Monthly. 69 (1962) p.4–8